

# The Černý conjecture for small automata: experimental report

Jakub Kowalski, Marek Szykuła

Department of Mathematics and Computer Science, University of Wrocław  
`{kot,msz}@ii.uni.wroc.pl`

**Abstract.** We present a report from a series of experiments involving computation of the shortest reset words for automata with small number of states. We confirm that the Černý conjecture is true for all automata with at most 11 states on 2 letters. Also some new interesting results were obtained, including the third gap in the distribution of the shortest reset words and new slowly synchronizing classes of automata.

**Keywords:** Černý conjecture, synchronizing word, nonisomorphic automata

## 1 Introduction

We deal with complete deterministic finite automata  $A = \langle Q, \Sigma, \delta \rangle$  with the state set  $Q$ , the input alphabet  $\Sigma$ , and the transition function  $\delta : Q \times \Sigma \rightarrow Q$ . The action of  $\Sigma$  on  $Q$  given by  $\delta$  is denoted simply by concatenation:  $\delta(q, a) = qa$ . This action extends naturally to the action  $qw$  of words for any  $w \in \Sigma^*$ . If  $|Qw| = 1$ , that is, the image of  $Q$  by  $w$  consists of a single state, then  $w$  is called a *reset* (or *synchronizing*) word for  $A$ , and  $A$  itself is called *synchronizing*. The length of the shortest reset word of an automaton is called the *reset length*. An automaton is *irreducible synchronizing* if it is synchronizing, and the automata obtained from it by removing any letter are not synchronizing. In this paper we deal with the *unlabeled model* of automata, in which the states and the letters are unlabeled. We say that two automata are isomorphic if there is a bijection between their states and a bijection between the letters, in this way isomorphic automata have the same reset length.

The Černý conjecture states that every synchronizing automaton  $A$  with  $n$  states has a reset word of length  $\leq (n - 1)^2$ . This conjecture was formulated by Černý in 1964, and is considered the most longstanding open problem in the combinatorial theory of finite automata. So far, the conjecture has been proved only for a few special classes of automata and a general cubic upper bound has been established (see Volkov [12] for general motivation and an excellent survey of the results, and Trahtman [11] for a recently found new cubic bound).

The reset lengths of automata are intensively studied recently with computer aid. There are experiments ([7,8]) showing that the expected reset length of a random automaton with  $n$  states is small and this is conjectured to be below linear bound. Another conjecture stated up by Peter Cameron during BCC conference in Exeter 2011, and in [8] states up that the number of synchronizing labeled automata to the number of all labeled automata tends to 1 as  $n \rightarrow \infty$ . Those conjectures seem also to be true in the unlabeled model, which we discuss in the experimental section based on our results. There are dedicated computational packages implementing different algorithms to study various aspects of automata synchronization, TESTAS [9] and recently developed COMPAS [4].

The Černý conjecture has been verified so far for all strongly connected automata with  $k = 2$  letters and  $n \leq 10$  states, and  $k \leq 4$  letters and  $n \leq 7$  states, and  $k = 3$  letters and  $n = 8$

states [10,11]. See also [1] for  $n = 9$  states. It is sufficient to check the conjecture only for strongly connected automata. All the results were obtained with computer aid. The results of the currently performed experiments show up also interesting trends in distribution of the reset length, especially for automata with the reset length close to the upper bound  $(n-1)^2$ . Those are called *slowly synchronizing automata*. For  $6 \leq n \leq 10$  there was found a gap in the distribution of the reset length, which is an absence of automata with the reset length in an interval just below the upper bound  $(n-1)^2$  (first reported in [10]). Also for  $n = 9$  the second gap was found, and an island between the gaps with automata with the reset length in  $[n^2 - 3n + 2, n^2 - 3n + 4]$  [1]. However these results obtained so far considered only strongly connected automata and are incomplete for  $n \geq 10$ , since the second gap was not reported here. Also neither of the experiments did ever reach  $n = 11$  states so far.

We have designed and implemented a new efficient algorithm to generate automata with small number of states and to compute the reset length for them. The method itself will be described in detail elsewhere. It allowed us to extend the known results and to verify the Černý conjecture for all automata with  $n \leq 11$  states and  $k = 2$  letters. Also we have obtained complete distributions of the reset length for all automata  $n \leq 10$ . For  $n = 11$  we report that the third gap in the distribution appears. Here we discuss the results of the experiments.

## 2 Experiments

### 2.1 The number of nonisomorphic automata

We have computed the exact numbers, as well as the complete distribution of the reset length, of nonisomorphic automata on 2 letters up to 10 states. The results are showed in Table 2.1. The total number of automata was known [5] and we have received exactly the same numbers of automata. We present also the number of synchronizing automata, the number of strongly connected, and the number of synchronizing strongly connected automata. For the numbers of nonisomorphic strongly connected automata on 2 labeled letters see [6], which are about 2 times larger than with unlabeled letters as we have considered (for example, there are 658,885 automata for  $n = 6$ ).

We can see that the fraction of synchronizing automata to all automata grows, and we may conjecture that it tends to 1 as was conjectured for the labeled model (P. Cameron and [8]). This growth is more significant in strongly connected automata, the fraction of synchronizing here for  $n = 10$  is about 0.999.

**Table 1.** The exact numbers of nonisomorphic automata with  $n$  unlabeled states on 2 unlabeled letters in the classes of all automata, synchronizing, strongly connected, and strongly connected synchronizing. In the last column there is the fraction of the number of synchronizing automata to all automata.

$n$	Total	Synchronizing	Strongly connected	S. c. and synchronizing	Synch./Total
2	7	4	4	2	0.57
3	74	51	29	21	0.69
4	1,474	1,115	460	395	0.76
5	41,876	34,265	10,701	10,180	0.82
6	1,540,696	1,318,699	329,794	322,095	0.86
7	68,343,112	60,477,844	12,310,961	12,194,323	0.88
8	3,540,691,525	3,210,707,626	538,586,627	536,197,356	0.91
9	209,612,916,303	193,589,241,468	26,959,384,899	26,904,958,363	0.92
10	13,957,423,192,794	13,070,085,476,528	1,518,185,815,760	1,516,697,994,964	0.94

## 2.2 Distribution of the length of the shortest reset word

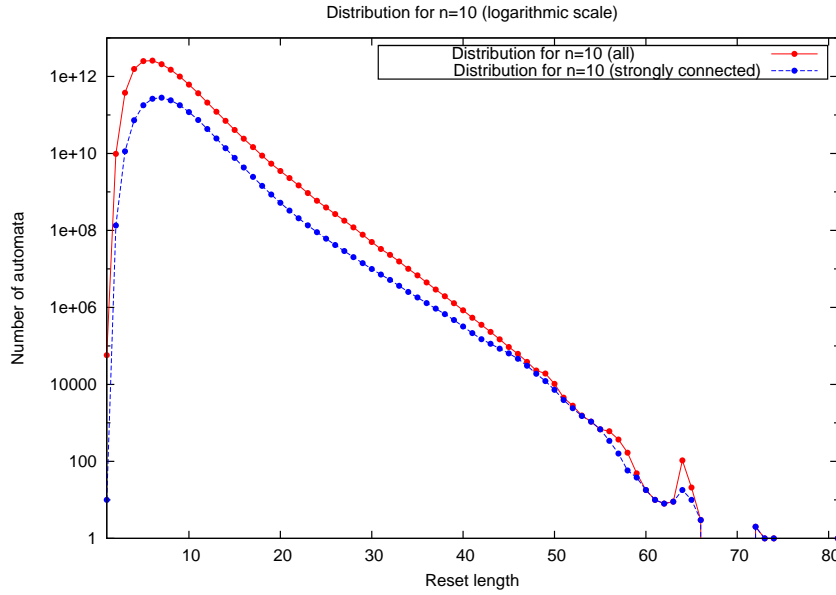
We have computed the complete distribution of the reset length of automata on 2 letters up to 10 states, and also we have verified the Černý conjecture for automata on 2 letters with 11 states. We confirm all the results reported in [10] and [1].

**The exact distributions for  $n \leq 10$ .** The paralleled computations have been performed on 16 computers with Intel(R) Core(TM) i7-2600 CPU 3.40GHz 4 cores and 1GB of RAM. Computing the complete distribution for all automata with  $n = 10$  states took above 800 days of total CPU time ( $\sim 13$  days of paralleled computations). Restricting to the class of strongly connected automata reduced this time to about 80 days of CPU ( $\sim 2$  days of paralleled computations).

Table 2.2 presents the obtained exact numbers of all and, separately, strongly connected nonisomorphic automata with  $n = 10$  states on  $k = 2$  letters with the longest reset lengths. The complete distribution for this case in the logarithmic scale is presented in Figure 2.2.

**Table 2.** The exact numbers  $N(\ell)$  of all and  $N_{sc}(\ell)$  of strongly connected nonisomorphic automata with 10 states on 2 letters with the shortest reset word of length  $\ell$ .

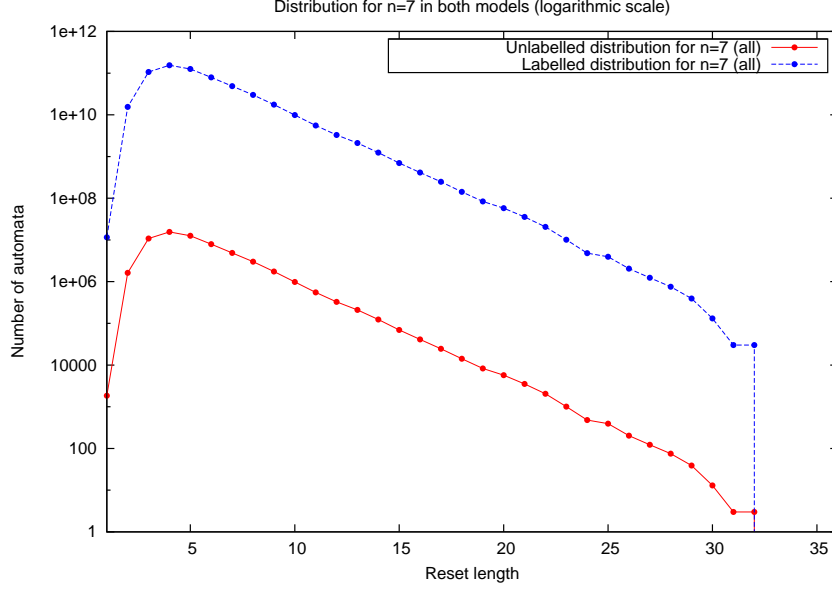
$\ell$	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81
$N(\ell)$	607	369	168	49	18	10	8	9	106	21	3	0	0	0	0	0	2	1	1	0	0	0	0	0	0	1
$N_{sc}(\ell)$	343	160	58	38	18	10	8	9	18	10	3	0	0	0	0	0	2	1	1	0	0	0	0	0	0	1



**Fig. 1.** The number of nonisomorphic automata for each reset length.

We were also able to compute the complete distribution in the labeled model for  $n \leq 7$ . In the labeled model the states and the letters are labeled. It may be interesting to compare regularities

of both the models for  $n = 7$  in Figure 2.2. We see that the shape of the both the distributions is very similar. This fact gives rise to suppose that statistical properties of the distribution in the labeled model (like the fraction of synchronizing ones etc.) are also present in the unlabeled model.



**Fig. 2.** The number of automata with 7 states on 2 letters in the both models for each reset length.

**The distribution for  $n = 11$ .** We have also verified the Černý conjecture for all automata with  $n = 11$  states on 2 letters. To do this it was sufficient to restrict the tested class of automata to strongly connected. However we have not obtained the complete distribution with due to an excessive number of automata. We have computed the exact number of automata only for long reset words. By skipping synchronizing letters at the beginning we were enumerating only irreducible synchronizing automata. The number of checked strongly connected automata in this case was 79,246,008,127,339. The total CPU time in this case was above 4 years ( $\sim 25$  days of paralleled computations).

Table 2.2 shows the obtained numbers of all nonisomorphic automata with  $n = 11$  states on  $k = 2$  letters with the longest reset lengths. We notice that the *third gap* have appeared for  $n = 11$  (for the report of the first gap see [10], and for the second [1]). The third gap does not occur in the distribution for  $n = 10$ . In the table there are presented some slowly synchronizing series corresponding to the reset lengths, which we describe in the next subsection.

We suppose that the irregularity in the upper part of distributions – with automata with long reset lengths – occur also for larger numbers of states and lead to more number of gaps for larger number of states, and we may state up the gap conjecture:

*Conjecture 1.* For any natural number  $g$ , there exists a big enough natural number  $n$  such that, in the distribution of the reset length of all nonisomorphic automata with  $n$  states on 2 letters there are at least  $g$  gaps.

**Table 3.** The numbers  $N_{\text{sc}}(\ell)$  of strongly connected nonisomorphic automata with 11 states on 2 letters with the reset length  $\ell$ .

$\ell$	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
$N_{\text{sc}}(\ell)$	$\leq 4$	$\leq 4$	$\leq 3$	0	0	9	22	12	2	1	0	0	0	0	0	3	2	1	0	0	0	0	0	0	0	1
Classes				–	–			$\mathcal{D}_n$	$\mathcal{W}''_n$	$\mathcal{B}_n$	–	–	–	–	–	$\mathcal{D}''_n$	$\mathcal{W}_n$	$\mathcal{D}'_n$	–	–	–	–	–	–	–	$\mathcal{C}_n$
								$\mathcal{C}_n^3$	$\mathcal{W}'_n$							$\mathcal{B}_n$	$\mathcal{C}_n^2$									
								$\dots$								$\mathcal{W}_n$										

### 2.3 Slowly synchronizing classes of automata

By extending the results in [1] to  $n = 11$  we have found some new slowly synchronizing classes of automata, but there is not any new in the island between the first and the second gap, all of them were found for  $n = 9$  states. We suppose so then that there are only these classes in the island for each  $n \geq 9$ :

*Conjecture 2.* In the distribution of the reset length for each number of states  $n \geq 9$  and 2 letters, there is an island  $[n^2 - 3n + 2, n^2 - 3n + 4]$  and there are exactly 4 automata  $\mathcal{D}'_n, \mathcal{W}_n, \mathcal{D}''_n, \mathcal{B}_n$  if  $n$  is even, and 2 more automata  $\mathcal{C}_n^2, \mathcal{W}'_n$  if  $n$  is odd.

In Table 2.2 we present some of the corresponding classes of automata in the distribution. Here is a survey of known slowly synchronizing classes with the longest reset length, with some new found classes.

- $\mathcal{C}_n$ :  $n^2 - 2n + 1$ , the Černý automaton [3].
- $\mathcal{D}'_n$ :  $n^2 - 3n + 4$  [1].
- $\mathcal{W}_n$ :  $n^2 - 3n + 3$  [1].
- $\mathcal{C}_n^2$ :  $n^2 - 3n + 3$  for odd  $n$ , defined below.
- $\mathcal{D}''_n$ :  $n^2 - 3n + 2$  [1].
- $\mathcal{W}'_n$ :  $n^2 - 3n + 2$  [1] (only mentioned there).
- $\mathcal{B}_n$ :  $n^2 - 3n + 2$  for odd  $n$  [2,1].
- $\mathcal{W}_n$ :  $n^2 - 3n + 2$  for odd  $n$ , defined below.
- $\mathcal{B}_n$ :  $n^2 - 4n + 7$  for odd  $n$ , defined below.
- $\mathcal{W}''_n$ :  $n^2 - 4n + 6$  [1] (only mentioned there).
- $\mathcal{W}''_n$ :  $n^2 - 4n + 6$  for odd  $n$ , defined below.
- $\mathcal{D}_n$ :  $n^2 - 4n + 5$  [2].
- $\mathcal{C}_n^3$ :  $n^2 - 4n + 5$  for odd  $n$ , defined below.

**The automaton  $\mathcal{W}'_n$ .** This automaton is similar to  $\mathcal{W}_n$  but with loops instead of a cycle on the second letter:

$$\delta(i, a) = \begin{cases} i+1 & \text{if } 1 \leq i \leq n-1, \\ 2 & \text{if } i = n; \end{cases} \quad \delta(i, b) = \begin{cases} 2 & \text{if } i = 1, \\ 1 & \text{if } i = 2, \\ i & \text{if } 3 \leq i \leq n-2, \\ n & \text{if } i = n-1. \end{cases}$$

**$\mathcal{C}_n^s$  automata.** We define here an generalization of the Černý automaton  $\mathcal{C}_n$ , by parametrized the distance between the states joined by the edge on the second letter.

**Definition 1.** The automaton  $\mathcal{C}_n^s$  for  $s < n$  is defined as follows:

$$\delta(i, a) = \begin{cases} i+1 & \text{if } 1 \leq i \leq n-1, \\ 1 & \text{if } i = n; \end{cases} \quad \delta(i, b) = \begin{cases} i & \text{if } 2 \leq i \leq n, \\ s & \text{if } i = 1. \end{cases}$$

Clearly  $\mathcal{C}_n^1 = \mathcal{C}_n$ . The automata for  $s > 1$  are synchronizing only for odd  $n$  and their reset length is  $(n-1)^2 - (n-2)(s-1)$ .

**The automaton  $\dot{\mathcal{B}}_n$ .** This automaton is similar to  $\mathcal{B}_n$ , the first letter forms a complete circle but the second crossing edge in the second letter is placed one state later than the first crossing edge.

$$\delta(i, a) = \begin{cases} i+1 & \text{if } 1 \leq i \leq n-1, \\ 1 & \text{if } i = n; \end{cases} \quad \delta(i, b) = \begin{cases} 2 & \text{if } i = 1, \\ 5 & \text{if } i = 2, \\ i+1 & \text{if } 3 \leq i \leq n-1, \\ 3 & \text{if } i = n-1. \end{cases}$$

The automata are synchronizing only for odd  $n$  and their reset length is  $n^2 - 4n + 7$ .

**The automata  $\mathcal{W}''_n$  and  $\dot{\mathcal{W}}''_n$ .** The automaton  $\mathcal{W}''_n$ :

$$\delta(i, a) = \begin{cases} i+1 & \text{if } 1 \leq i \leq n-1, \\ 2 & \text{if } i = n; \end{cases} \quad \delta(i, b) = \begin{cases} i & \text{if } 2 \leq i \leq n-3 \text{ and } i = n-1 \\ 1 & \text{if } i = n-2, \\ n-2 & \text{if } i = 1. \end{cases}$$

The automaton  $\dot{\mathcal{W}}''_n$ :

$$\delta(i, a) = \begin{cases} i+1 & \text{if } 1 \leq i \leq n-1, \\ 2 & \text{if } i = n; \end{cases} \quad \delta(i, b) = \begin{cases} i & \text{if } 2 \leq i \leq n-3 \\ 2 & \text{if } i = n-1, \\ 1 & \text{if } i = n-2, \\ n-2 & \text{if } i = 1. \end{cases}$$

The automata are synchronizing only for odd  $n$  and their reset length is  $n^2 - 4n + 6$ .

## References

1. D. Ananichev, V. Gusev, and M. Volkov. Slowly synchronizing automata and digraphs. In *Mathematical Foundations of Computer Science 2010*, volume 6281 of *LNCS*, pages 55–65. 2010.
2. D. Ananichev, M. Volkov, and Y. Zaks. Synchronizing automata with a letter of deficiency 2. In *Developments in Language Theory*, volume 4036 of *LNCS*, pages 433–442. 2006.
3. J. Černý. Poznámka k homogénnym eksperimentom s konečnými automatami. *Matematicko-fyzikálny Časopis Slovenskej Akadémie Vied*, 14(3):208–216, 1964. In Slovak.
4. K. Chmiel and A. Roman. COMPAS - A computing package for synchronization. In *Implementation and Application of Automata*, volume 6482 of *LNCS*, pages 79–86. 2011.
5. M. Harrison. A census of finite automata. *Canadian journal of mathematics*, 17:100–113, 1965.
6. V. A. Liskovets. Enumeration of non-isomorphic strongly connected automata. *Vesci Akad. Navuk BSSR Ser. Fiz.-Téhn. Navuk*, (3):26–30, 1971. In Russian.
7. A. Roman. Genetic algorithm for synchronization. In *Language and Automata Theory and Applications*, volume 5457 of *LNCS*, pages 684–695. 2009.
8. E. Skvortsov and E. Tipikin. Experimental study of the shortest reset word of random automata. In *Implementation and Application of Automata*, volume 6807 of *LNCS*, pages 290–298. 2011.

9. A. N. Trahtman. A package TESTAS for checking some kinds of testability. In *Implementation and Application of Automata*, volume 2608 of *LNCS*, pages 228–232. 2003.
10. A. N. Trahtman. An efficient algorithm finds noticeable trends and examples concerning the Černý conjecture. In *Mathematical Foundations of Computer Science*, volume 4162 of *LNCS*, pages 789–800. 2006.
11. A. N. Trahtman. Modifying the upper bound on the length of minimal synchronizing word. In *Fundamentals of Computation Theory*, volume 6914 of *LNCS*, pages 173–180. 2011.
12. M. Volkov. Synchronizing automata and the Černý conjecture. In *Language and Automata Theory and Applications*, volume 5196 of *LNCS*, pages 11–27. 2008.